Available online at www.sciencedirect.com

ScienceDirect

Physics Procedia 68 (2015) 115 – 119

Physics Procedia

28th Annual CSP Workshops on "Recent Developments in Computer Simulation Studies in Condensed Matter Physics", CSP 2015

Teraflops and beyond: GPU-based MD studies of emergent phenomena

D.C. Rapaport[a]

[a]*Department of Physics, Bar-Ilan University, Ramat-Gan, Israel 52900*

Abstract

Molecular dynamics simulations of emergent phenomena are often computationally demanding because of the broad range of length and time scales that must be covered, ranging from the individual particles, out to where the collective behavior is expressed; the fact that simulations of this type are often subject to unpredictable outcomes is a further complication. The computations required for these studies benefit substantially from massively parallel GPU-based implementation, with even a single GPU typically providing an order of magnitude performance gain over a conventional CPU. A sampling of recently obtained exploratory results involving atomistic hydrodynamics, granular segregation and self-assembly are discussed, along with key aspects of the methodology; the rich behavior observed provides a hint of the kinds of phenomena that can be explored, and what might be achieved given adequate resources.

Peer-review under responsibility of The Organizing Committee of CSP 2015 Conference

Keywords:
emergent phenomena, molecular dynamics simulation, GPU computation, atomistic hydrodynamics, granular segregation, self-assembly

PACS: 02.70.Ns, 45.70.Mg, 47.20.Qr, 47.55.pb, 81.16.Fg

1. Introduction

This paper deals with emergent phenomena, in particular, the kinds of results that can be obtained from molecular dynamics (MD) simulation, an approach that employs a detailed particle-level description of the system. In the examples considered here, the particles themselves correspond to atoms or small molecules, to macroscopic grains, or to large molecular complexes, and the emergent character of the behavior is reflected in complex coherent flows, spatial segregation based on particle properties, or self-assembled structures of specific types. While complex behavior of this kind is frequently encountered in the real world, there can be no question that its appearance in computer simulations of highly simplified models is in itself surprising. The simulations described here require relative large and lengthy computations; GPU-based (graphic processing unit) computing is ideally suited for MD work once the necessary algorithmic adaptations have been made.

E-mail address: rapaport@mail.biu.ac.il

Peer-review under responsibility of The Organizing Committee of CSP 2015 Conference
doi:10.1016/j.phpro.2015.07.119

2. Methodology

MD simulation (Rapaport, 2004) is a well-established approach used in a broad range of physical, biophysical, materials-science and engineering problems. The common element is the numerical solution of a classical many-body problem, but the details are problem specific. Many of the applications involve systems in thermal equilibrium, but this is not a requirement, and indeed the examples considered here are far from equilibrium, and in most cases are also externally driven. Thus there is very little theoretical guidance as to the expected outcomes, especially in view of the limited length and time scales accessible to MD. The models employed are about as simple as possible, but the simulations themselves are carried out in full detail to avoid any extraneous effects of even further simplification, such as the combining of discrete and continuum descriptions when dealing with fluids and granular media, or using spheres rather than customized particles when studying self-assembly.

The algorithms for MD simulation must be modified to accommodate the limitations of GPU architecture. These issues are addressed using an approach devised originally for vector supercomputers, namely methods for systematically accessing what amounts to randomly organized particle data. Since neither vector processors nor GPUs were created with MD in mind the overall performance will be far from the the theoretical maximum; nevertheless, algorithms can be devised that achieve respectable performance with a major improvement over conventional scalar CPUs, but at the price of program complexity.

The main change to the overall MD approach concerns the way the neighbor list is constructed. In the conventional approach, the neighbor list enumerates particle pairs that are inside or close to the interaction cutoff and is constructed by first assigning particles to cells that are described using linked lists. These tasks are totally unsuited to the GPU approach because of their heavy reliance on efficient random memory access, a feature totally absent from the GPU design. The GPU approach requires the following changes: The revised neighbor list references each particle pair twice to allow the force acting on each particle to be accumulated independently. Cell occupancy is represented by a set of layers (effectively a matrix) with the nth layer recording the identity of the nth particle in the cell, so that the total number of layers equal to the maximum cell occupancy. Finally, building the neighbor list requires looping over all possible layer pairs, with the contents of adjacent cells now being examined in parallel for possibly interacting particles. A more detailed description of the GPU algorithms appears in (Rapaport, 2011).

Both the neighbor list construction and the force evaluation access particle data in a totally unordered manner. Two steps are taken to avoid serious performance degradation. The first is to explicitly read data using one of the hardware caches; this requires only minor program changes. However, to ensure that caching can benefit from data localization, it is also necessary to rearrange the particle data so that particles that are in physical proximity are also stored in nearby memory locations (or at least in a few small clusters in memory); this sorting operation is easily done at regular intervals using available cell membership information and can have a major (5X) impact on performance.

These and other issues must be taken into account during algorithm design and programming, but since they are common to different kinds of MD simulation the techniques can be reused without change. Further specific extensions required in the different simulations include rotational motion of extended particles, multiple particle sizes, different kinds of hard-wall boundaries, dissipative interactions for granular media, and thermostats, all of which are readily implemented.

Performance tests show the resultant major improvements. The GPU code uses single precision floating-point arithmetic, which is adequate for this type of MD; double precision code would typically run at less than half the speed. Comparing the Nvidia K20 GPU (with 2496 cores) and a single core of an Intel Xeon E5-2637v2 (running at 3.5GHz) for soft-sphere and LJ systems at moderately high densities and with several hundred thousand to millions of particles shows gains of 23X and 47X respectively. Even if all the eight cores available on a dual-CPU machine are used for the test by means of multithreading or MPI the performance gain is still 4-8X; note that in both these cases algorithm and program changes are required, although less drastic than for the GPU. Thus the GPU is a clear winner in tests of this kind. The results described below were obtained using a dual-GPU workstation, with the GPUs running independently on different computations.

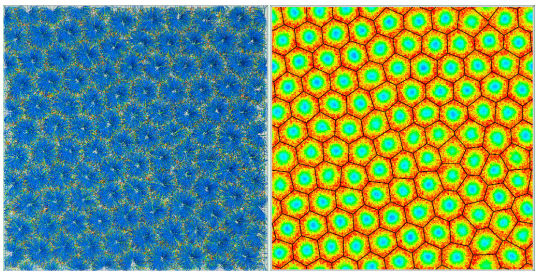

Fig. 1. Flow and temperature fields showing the final state of the Rayleigh–Bénard simulation.

3. Results

The systems discussed in this section have been considered previously, but GPU availability has allowed considerably more extensive exploration of these problems. The discussion that follows outlines the key features of the models and demonstrates the kinds of behavior that can be encountered. More detailed descriptions appear in a recent review (Rapaport, 2014) which also contains a full bibliography.

3.1. Hydrodynamic instability

The atomistic approach to fluid dynamics explores the behavior of fluids modeled using discrete particles, as opposed to the normal continuum description that is completely oblivious of the atomistic nature of matter. For obvious reasons it cannot compete with continuum numerical techniques, and so the initial step is to establish which aspects of hydrodynamics persist down to the very small length and time scales accessible to MD. If qualitatively reasonable behavior can be obtained then this should be followed by quantitative analysis.

The two cases considered here are the Rayleigh–Bénard (RB) and Taylor–Couette (TC) instabilities. The RB instability occurs in a fluid layer heated from below when the Rayleigh number, a dimensionless quantity dependent on temperature difference and layer height, exceeds a critical value, and which manifests itself by the appearance of convective flow cells. The TC instability occurs in an annular layer confined between two cylinders, the inner of which rotates, and is marked by toroidal rolls superimposed on the sheared annular flow appearing when the Taylor number, a quantity dependent on flow velocity and layer width, exceeds a critical value. Theoretical estimates for these critical values are known, and are achievable in MD simulations if parameters such as the temperature difference or flow velocity are sufficiently large. Both systems include nonslip boundaries, in the RB case these are the upper and lower horizontal walls while the lateral boundaries are periodic, and in the TC case the inner and outer cylinder walls are rough and the end caps smooth. The boundaries also participate in heat transfer, in the RB case by injecting or absorbing heat to maintain a thermal gradient, and in the TC case by removing the heat generated by the shear flow.

Fig. 1 shows the outcome of a simulation of an RB system with 1.08×10^7 atoms over 1.8×10^7 timesteps, in a container whose ratio of length to height is 20:1. Coarse-grained space and time averaging is used to extract the effective fluid behavior from the MD particle trajectories. The final flow is shown in the first panel, while the second shows the temperature variation in a cross-section at approximately a quarter of the height from the base, with superimposed convection cell boundaries determined by automated Voronoi analysis. The overall behavior is an array of convection cells, mainly regular, but with some defects, where cooler fluid descends at the center of each cell and hot fluid ascends at the periphery. The characteristic wavelength of the cell pattern is consistent with theory.

Fig. 2 shows early and final states of a TC system containing 2.35×10^6 atoms in a narrow annular region whose length to width ratio is 16:1. To reduce computational effort, only one quadrant of the system is considered, with special periodic boundaries used to allow for the rest of the container. The azimuthal shear flow is subtracted from coarse-grained flow averages and an overall azimuthal average of the axial and radial flow components is shown, with color coding used to distinguish counter-rotating toroidal rolls. Between the initial appearance of the rolls and the final state after 1.9×10^7 timesteps there is some roll merging and repositioning, and the overall wavelength of the final roll array is close to what is expected theoretically.

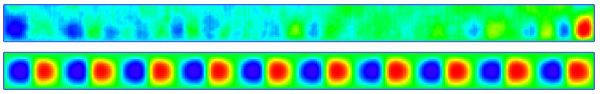

Fig. 2. Mean toroidal flow in the Taylor–Couette system early in the run and at the end.

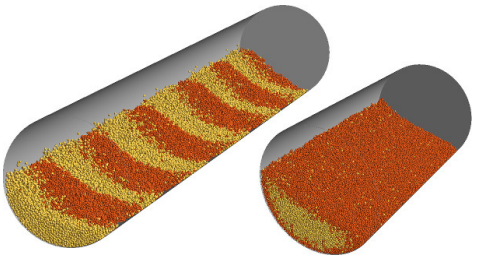

Fig. 3. Axial and radial segregation in a rotating drum.

3.2. Granular segregation

Granular media can be modeled using normal MD methods augmented by friction and damping forces, with the particles in the MD simulation corresponding to actual macroscopic grains. Such systems must be externally driven to ensure that the system does not come to rest, but the fact that energy is not conserved is of no concern to the MD approach.

Designing the granular particles involves deciding on particle shape and specifying the interparticle forces. Particle-wall interactions must also be specified, although these are often based on the interparticle forces. The grains are typically represented as spheres, although extended particles can also be used. Indeed, in the present study of segregation in a rotating drum, each grain is represented by four spheres fused into a rigid tetrahedral structure; the spacing between the spheres determines the degree to which these particles can resist sliding by interlocking. The interactions include a combination of normal (to the surface at the point of impact) excluded-volume and damping forces, together with optional tangential forces that represent the effects of sliding and static friction. The last of these interactions is more difficult to represent since it typically requires keeping track of the particles while in contact; in the drum problem the extended particles accomplish the same effect, while in the sawtooth problem it is not needed.

Fig. 3 shows examples of axial and radial segregation in a rotating drum for systems containing 10^5 and 1.35×10^5 particles respectively, in the former case after 1500 revolutions (10^7 timesteps) and 650 in the latter. The fill level in the run with radial segregation is higher than for the axial segregation case, and the drum is shorter to reduce the particle count. The initial state is a uniform mixture of the two kinds of particles that differ in size and surface roughness, reflected both in the spacing of the spheres forming the grains and the sliding friction coefficient. There is also transient radial segregation in the former case (not shown here), although its effect is less dramatic. In both examples the states shown appear to be stable.

Fig. 4 shows the outcome of simulations of granular mixtures on a vibrating sawtooth base, where the sawtooth profiles are asymmetric. In the case of the rectangular container the sawtooth grooves are linear, while the base of the cylindrical container has concentric circular grooves. The grains are spheres of two sizes, but are otherwise identical, and the systems contain 2×10^5 and 3.1×10^5 particles respectively. After a sufficiently large number of vibration cycles (9×10^4 and 4×10^4) the two grain species can be seen to exhibit relatively strong segregation.

3.3. Supramolecular self-assembly

The spontaneous self-assembly of molecules into large-scale structures is a very active field of study, with an enormous variety of different scenarios. The present study addresses just one of these problems, namely the formation

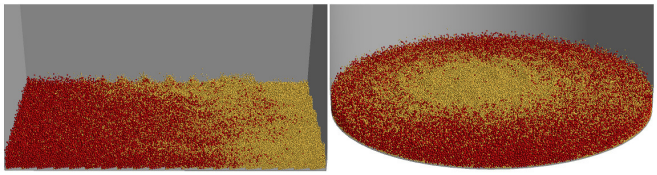

Fig. 4. Segregation in granular systems with a vibrating sawtooth base.

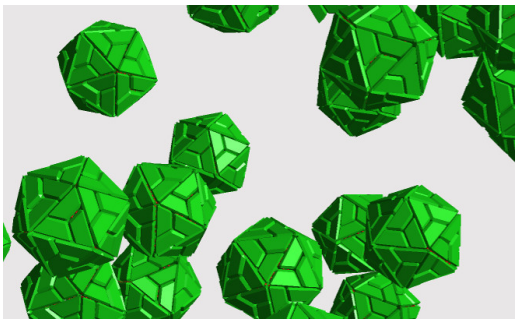

Fig. 5. Some of the complete shells in the self-assembly simulation.

of polyhedral shells from a set of identical units, a process motivated by the in vitro growth of viral capsid shells, and in particular, 60-faced T=1 shells built from trapezoidal monomers.

The self-assembling particles themselves are formed from multiple spheres fused into an array with an overall trapezoidal pyramid shape, whose dimensions and angles are consistent with the desired structure. The lateral faces also include attraction sites through which the bonding forces act; particles are considered bonded when they are suitably aligned with all of the multiple bonding sites in the corresponding faces lying within a specified distance much less than the attraction range. Bonds are readily broken, especially for particles attached to a cluster by just a single bond; this reversibility turns out to be important in achieving a high yield of defect-free complete shells. Particles are immersed in a soft-sphere solvent; the presence of an explicit rather than an implicit (or no) solvent helps ensure that the overall dynamics of the monomers and the partial and complete assemblies is correctly characterized.

Fig. 5 shows a close-up view of part of the final state in a simulation containing 4752 extended particles, potentially able to form 79 shells, with the remainder of the 2.16×10^5-particle system consisting of solvent atoms. The solvent atoms are not shown because they would completely obscure the interior, and monomers are also removed. A total of 57 complete shells have succeeded in forming (a 72% yield) after 3.2×10^8 timesteps, with 3 nearly-complete shells, while all the remaining particles are unbonded; the surprising aspect of this result is the very strong bimodal nature of the size distribution. The behavior is very sensitive to the strength of the attraction between particles (or, equivalently, the inverse temperature); a slightly reduced value results in no shell growth and just small clusters, while a slight increase leads to excessively rapid growth with a broad distribution of large clusters, none of them complete shells, and total monomer depletion.

References

Rapaport, D.C., 2004. The Art of Molecular Dynamics Simulation. 2nd ed., Cambridge University Press, Cambridge.
Rapaport, D.C., 2011. Enhanced molecular dynamics performance with a programmable graphics processor. Computer Phys. Comm. 182, 926.
Rapaport, D.C., 2014. Molecular dynamics simulation: a tool for exploration and discovery using simple models. J. Phys.: Condens. Matter 26, 503104.